(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2019 年 4 月 18 日 (18.04.2019)

WIPO | PCT

(10) 国际公布号
WO 2019/072091 A1

(51) 国际专利分类号:
G06K 9/62 (2006.01) *G06F 17/30* (2006.01)

(21) 国际申请号: PCT/CN2018/107969

(22) 国际申请日: 2018 年 9 月 27 日 (27.09.2018)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710948881.3 2017年10月12日 (12.10.2017) CN

(71) 申请人: 北京京东尚科信息技术有限公司 (**BEIJING JINGDONG SHANGKE INFORMATION TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市海淀区杏石口路65号西杉创意园四区11号楼东段1-4层西段1-4层, Beijing 100195 (CN)。
北京京东世纪贸易有限公司(**BEIJING JINGDONG CENTURY TRADING CO., LTD.**) [CN/CN]; 中国北京市大兴区北京经济技术开发区科创十一街18号C座2层201室, Beijing 100176 (CN)。

(72) 发明人: 余星梅(**YU, Xingmei**); 中国北京市海淀区杏石口路65号西杉创意园四区11号楼东段1-4层西段1-4层, 由北京京东尚科信息技术有限公司转交, Beijing 100195 (CN)。 陈海勇(**CHEN, Haiyong**); 中国北京市海淀区杏石口路65号西杉创意园四区11号楼东段1-4层西段1-4层, 由北京京东尚科信息技术有限公司转交, Beijing 100195 (CN)。 邵佳帅(**SHAO, Jiashuai**); 中国北京市海淀区杏石口路65号西杉创意园四区11号楼东段1-4层西段1-4层, 由北京京东尚科信息技术有限公司转交, Beijing 100195 (CN)。

(74) 代理人: 北京律智知识产权代理有限公司 (**BEIJING INTELLEGAL INTELLECTUAL PROPERTY AGENT LTD.**); 中国北京市朝阳区慧忠路5号B1605、B1606、B1607, Beijing 100101 (CN)。

(81) 指定国(除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,

(54) **Title:** METHOD AND APPARATUS FOR USE IN DETERMINING TAGS OF INTEREST TO USER

(54) 发明名称: 用于确定用户兴趣标签的方法及装置

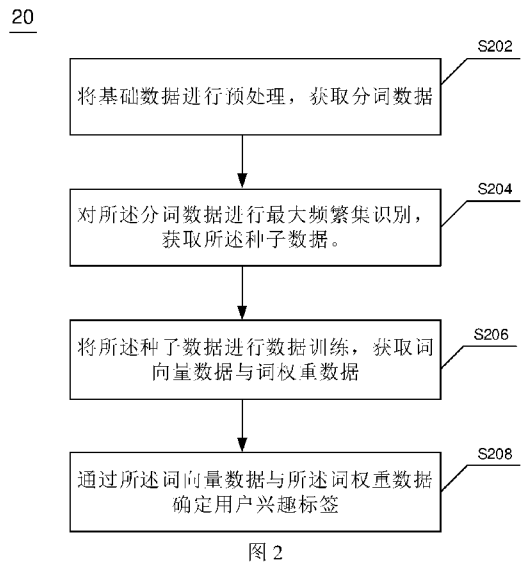

图 2

S202 Pre-process basic data to obtain word segmentation data
S204 Perform maximal frequent set recognition on the word segmentation data to obtain seed data
S206 Perform data training on the seed data to obtain word vector data and word weighting data
S208 Determine tags of interest to a user by means of the word vector data and the word weighting data

(57) **Abstract:** Disclosed in the present application are a method and apparatus for use in determining tags of interest to a user, relating to the field of computer information processing, wherein the method comprises: pre-processing basic data to obtain word segmentation data; performing maximal frequent set recognition on the word segmentation data to obtain seed data; performing data training on the seed data to obtain word vector data and word weighting data; and determining tags of interest to the user by means of the word vector data and the word weighting data. The method and apparatus disclosed by the present application that are used for determining tags of interest to a user may effectively determine subjects of interest to the user, reducing time spent on manual processing.

(57) 摘要: 本申请公开一种用于确定用户兴趣标签的方法及装置。涉及计算机信息处理领域，该方法包括：将基础数据进行预处理，获取分词数据；对所述分词数据进行最大频繁集识别，获取种子数据；将所述种子数据进行数据训练，获取词向量数据与词权重数据；以及通过所述词向量数据与所述词权重数据确定用户兴趣标签。本申请公开的用于确定用户兴趣标签的方法及装置，能够有效的确定用户的兴趣主题，减少人工处理时间。

WO 2019/072091 A1

[见续页]

GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) **指定国**(除另有指明，要求每一种可提供的地区保护)：ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

用于确定用户兴趣标签的方法及装置

技术领域

本发明涉及计算机信息处理领域，具体而言，涉及一种用于确定用户兴趣标签的方法及装置。

背景技术

随着网络购物的普及推广，购物网站之间的竞争越发激烈，电商崛起，企业要想长期稳定的生存，首先必须吸引用户，其次需要经营用户，从而使得用户成为企业的忠诚用户。如何很好的经营用户，是一个难题，随着用户行为数据的记录，数据挖掘算法技术的成熟，企业可以通过多种方法来经营用户，如何将用户感兴趣的东西推送给用户，在电子商务中异常重要。在这个过程中，识别用户兴趣是非常重要的一环。基于对用户的兴趣的识别，其中最为常见也最核心的就是对用户进行精准营销，在对的时间把对的商品推荐给对的人。要对用户进行精准营销，亦或是某供应商需要把自己的商品卖给对的人，就需要借助用户画像来实现，而用户兴趣度标签是确定用户对某个品类或者品牌想要购买的一个兴趣程度，即企业可以根据用户的兴趣标签推荐合适的商品给用户，供应商可以根据兴趣标签圈定对自己商品感兴趣的人群进行营销，从而企业/供应商以及用户达到双赢。

用户兴趣多种多样，在不同的行业，需要关注的用户兴趣不同，电商行业关注的是影响用户购买的兴趣爱好。所以，目前一般的思路是直接对用户在网站购买或者浏览过的商品使用 LDA 主题模型，得到若干兴趣主题，然后人工对这部分兴趣主题进行标注。直接使用 LDA 主题模型得到的结果重复率高，有效性较低，后期需要的人工标注和过滤的工作量很大。

因此，需要一种新的用于确定用户兴趣标签的方法及装置。

在所述背景技术部分公开的上述信息仅用于加强对本发明的背景的理解，因此它可以包括不构成对本领域普通技术人员已知的现有技术的信息。

发明内容

有鉴于此，本发明提供一种用于确定用户兴趣标签的方法及装置，能够有效的确定用户的兴趣主题，减少人工处理时间。

本发明的其他特性和优点将通过下面的详细描述变得显然，或部分地通过本发明的实践而习得。

根据本发明的一方面，提出一种用于确定用户兴趣标签的方法，该方法包括：将基础数据进行预处理，获取分词数据；对所述分词数据进行最大频繁集识别，获取种子数据；将所述种子数据进行数据训练，获取词向量数据与词权重数据；以及通过所述词向量数据与所述词权重数据确定用户兴趣标签。

在本公开的一种示例性实施例中，所述将基础数据进行预处理，获取分词数据，包括：通过用户历史购物数据生成所述基础数据；以及对所述基础数据进行分词处理，生成所述分词数据。

在本公开的一种示例性实施例中，所述对所述分词数据进行最大频繁集识别，获取种子数据，包括：根据预定条件，获取所述分词数据中所有的组合数据；对每一种组合数据，根据其订单数量，确定所述组合数据的频繁集；对所述频繁集进行最大频繁集计算，获取种子数据。

在本公开的一种示例性实施例中，所述对所述分词数据进行最大频繁集识别，获取种子数据，包括：通过数据仓库的分布式计算架构，对所述分词数据进行最大频繁集识别，获取所述种子数据。

在本公开的一种示例性实施例中，所述将所述种子数据进行数据训练，包括：通过三层贝叶斯模型对所述种子数据进行数据训练。

在本公开的一种示例性实施例中，还包括：通过历史数据，获取用户购买数据，所述购买数据包括购买产品次数以及购买产品标识。

在本公开的一种示例性实施例中，所述通过所述词向量数据与所述词权重数据确定用户的兴趣标签，包括：通过所述用户购买数据，确定所述用户的词向量数据以及词权重数据；通过所述用户的词向量数据以及词权重数据，计算所述用户的兴趣数值；通过所述兴趣数值确定所述用户的所述兴趣标签。

在本公开的一种示例性实施例中，所述通过所述用户的词向量数据以及词权重数据，计算所述用户的兴趣数值，包括：

$Sum=(a*Q)$;其中，Sum 为用户的所述兴趣数值，a 为用户购买产品次数，Q 为产品对应的词权重。

在本公开的一种示例性实施例中，所述通过所述兴趣数值确定所述用户的所述兴趣标签，还包括：判断所述兴趣数值是否大于预定阈值；以及将大于预定阈值的所述兴趣数值对应的兴趣标签确定为所述用户的兴趣标签。

在本公开的一种示例性实施例中，还包括：通过所述用户的所述兴趣标签进行信息推广。

根据本发明的一方面，提出一种用于确定用户兴趣标签的装置，该装置包括：基础模块，用于将基础数据进行预处理，获取分词数据；种子模块，用于对所述分词数据进行最大频繁集识别，获取种子数据；训练模块，用于将所述种子数据进行数据训练，获取词向量数据与词权重数据；以及标签模块，用于通过所述词向量数据与所述词权重数据确定用户兴趣标签。

根据本发明的一方面，提出一种电子设备，该电子设备包括：一个或多个处理器；存储装置，用于存储一个或多个程序；当一个或多个程序被一个或多个处理器执行，使得一个或多个处理器实现如上文的方法。

根据本发明的一方面，提出一种计算机可读介质，其上存储有计算机程序，其特征在于，程序被处理器执行时实现如上文中的方法。

根据本发明的用于确定用户兴趣标签的方法及装置，能够有效的确定用户的兴趣主题，减少人工处理时间。

应当理解的是，以上的一般描述和后文的细节描述仅是示例性的，并不能限制本发明。

附图说明

图 1 是根据一示例性实施例示出的一种用于确定用户兴趣标签的方法的系统架构。

图 2 是根据一示例性实施例示出的一种用于确定用户兴趣标签的方法的流程图。

图 3 是根据一示例性实施例示出的一种用于确定用户兴趣标签的方法的示意图。

图 4 是根据另一示例性实施例示出的一种用于确定用户兴趣标签的方法的示意图。

图 5 是根据另一示例性实施例示出的一种用于确定用户兴趣标签的方法的流程图。

图 6 是根据一示例性实施例示出的一种用于确定用户兴趣标签的方法的示意图。

图 7 是根据另一示例性实施例示出的一种用于确定用户兴趣标签的方法的示意图。

图 8 是根据一示例性实施例示出的一种用于确定用户兴趣标签的方法的示意图。

图 9 是根据另一示例性实施例示出的一种用于确定用户兴趣标签的方法的示意图。

图 10 是根据另一示例性实施例示出的一种用于确定用户兴趣标签的方法的流程图。

图 11 是根据一示例性实施例示出的一种用于确定用户兴趣标签的装置的框图。

图 12 是根据一示例性实施例示出的一种电子设备的框图。

图 13 是根据一示例性实施例示出的一种计算机可读介质示意图。

具体实施方式

现在将参考附图更全面地描述示例实施例。然而，示例实施例能够以多种形式实施，且不应被理解为限于在此阐述的实施例；相反，提供这些实施例使得本发明将全面和完整，并将示例实施例的构思全面地传达给本领域的技术人员。在图中相同的附图标记表示相同或类似的部分，因而将省略对它们的重复描述。

此外，所描述的特征、结构或特性可以以任何合适的方式结合在一个或更多实施例中。在下面的描述中，提供许多具体细节从而给出对本发明的实施例的充分理解。然而，本领域技术人员将意识到，可以实践本发明的技术方案而没有特定细节中的一个或更多，或者可以采用其它的方法、组元、装置、步骤等。在其它情况下，不详细示出或描述公知方法、装置、实现或者操作以避免模糊本发明的各方面。

附图中所示的方框图仅仅是功能实体，不一定必须与物理上独立的实体相对应。即，可以采用软件形式来实现这些功能实体，或在一个或多个硬件模块或集成电路中实现这些功能实体，或在不同网络和/或处理器装置和/或微控制器装置中实现这些功能实体。

附图中所示的流程图仅是示例性说明，不是必须包括所有的内容和操作/步骤，也不

是必须按所描述的顺序执行。例如，有的操作/步骤还可以分解，而有的操作/步骤可以合并或部分合并，因此实际执行的顺序有可能根据实际情况改变。

应理解，虽然本文中可能使用术语第一、第二、第三等来描述各种组件，但这些组件不应受这些术语限制。这些术语乃用以区分一组件与另一组件。因此，下文论述的第一组件可称为第二组件而不偏离本公开概念的教示。如本文中所使用，术语“及/或”包括相关联的列出项目中的任一个及一或多者的所有组合。

本领域技术人员可以理解，附图只是示例实施例的示意图，附图中的模块或流程并不一定是实施本发明所必须的，因此不能用于限制本发明的保护范围。

下面结合附图对本公开示例实施方式进行详细说明。

图 1 是根据一示例性实施例示出的一种用于确定用户兴趣标签的方法的系统架构。

如图 1 所示，系统架构 100 可以包括终端设备 101、102、103，网络 104 和服务器 105。网络 104 用以在终端设备 101、102、103 和服务器 105 之间提供通信链路的介质。网络 104 可以包括各种连接类型，例如有线、无线通信链路或者光纤电缆等等。

用户可以使用终端设备 101、102、103 通过网络 104 与服务器 105 交互，以接收或发送消息等。终端设备 101、102、103 上可以安装有各种通讯客户端应用，例如购物类应用、网页浏览器应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等。

终端设备 101、102、103 可以是具有显示屏并且支持网页浏览的各种电子设备，包括但不限于智能手机、平板电脑、膝上型便携计算机和台式计算机等等。

服务器 105 可以是提供各种服务的服务器，例如对用户利用终端设备 101、102、103 所浏览的购物类网站提供支持的后台管理服务器。后台管理服务器可以对接收到的产品信息查询请求等数据进行分析等处理，并将处理结果（例如推送信息、产品信息）反馈给终端设备。

需要说明的是，本申请实施例所提供的推广消息生成方法一般由服务器 105 执行，相应地，推送消息的展示网页一般设置于客户端 101 中。

应该理解，图 1 中的终端设备、网络和服务器的数目仅仅是示意性的。根据实现需要，可以具有任意数目的终端设备、网络和服务器。

图 2 是根据一示例性实施例示出的一种用于确定用户兴趣标签的方法的流程图。

如图 2 所示，在 S202 中，将基础数据进行预处理，获取分词数据。可例如，通过用户历史购物数据生成所述基础数据；以及对所述基础数据进行分词处理，生成所述分词数据。在现实场景中，用户在网站的一次或一段时间的购物行为都是围绕一定的目的或者兴趣爱好进行。在本实施例中，可例如假设用户每次下单是围绕某个兴趣进行，进而从数据仓库中提取所有用户一年的购物历史数据作为基础数据，基础数据可例如以(用户账号+订单+商品 id+商品名)为一行的形式存放。可例如，使用分词方法处理基础数据中商品的产品词，将同一个订单的产品词组合为一个产品词列表，产品词之间用逗号分割存储，此时的数据为分词数据，数据形式可例如为：订单+产品词列表的形式，基础数据格式与分

词数据可例如如图 3 所示。

在 S204 中，对所述分词数据进行最大频繁集识别，获取所述种子数据。项的集合称为项集。包含 k 个项的项集称为 k-项集，集合{computer,ativirus_software}是一个二项集。项集的出项频率是包含项集的事务数，简称为项集的频率，支持度计数或计数。注意，定义项集的支持度有时称为相对支持度，而出现的频率称为绝对支持度。如果项集 I 的相对支持度满足预定义的最小支持度阈值，则 I 是频繁项集。最大频繁集是指，如果频繁项集 L 的所有超集都是非频繁项集，那么称 L 为最大频繁项集或称最大频繁模式，记为 MFI (Maximal Frequent Itemset) 。频繁项集是最大频繁项集的子集。最大频繁项集中包含了频繁项集的频繁信息，且通常项集的规模要小几个数量级。所以在数据集中含有较长的频繁模式时挖掘最大频繁项集是非常有效的手段。可例如，通过数据仓库的分布式计算架构，对所述分词数据进行最大频繁集识别，获取所述种子数据。

在 S206 中，将所述种子数据进行数据训练，获取词向量数据与词权重数据。可例如，通过三层贝叶斯模型对所述种子数据进行数据训练。LDA（Latent Dirichlet Allocation）是一种文档主题生成模型，也称为一个三层贝叶斯概率模型，包含词、主题和文档三层结构。所谓生成模型，就是说，可认为一篇文章的每个词都是通过“以一定概率选择了某个主题，并从这个主题中以一定概率选择某个词语”这样一个过程得到。文档到主题服从多项式分布，主题到词服从多项式分布。通过 LDA 模型训练可例如获取种子数据中完整的词向量以及每个词的权重。

在 S208 中，通过所述词向量数据与所述词权重数据确定用户兴趣标签。对于每一个用户而言，均可以由词向量以及词权重计算获得该用户在某一分类下的所有的产品词及产品词权重。综合考虑该用户在某一分类下的所有的产品词及产品词权重（可例如为产品词与其对应的产品词权重乘积的形式），即可获得该用户的兴趣得分。可例如，判断所述兴趣数值是否大于预定阈值；以及将大于预定阈值的所述兴趣数值对应的兴趣标签确定为所述用户的兴趣标签。

根据本发明的用于确定用户兴趣标签的方法，通过对原始数据进行分词表示，进而采用三层贝叶斯网络对分词数据进行训练，获得词向量以及词权重，进而确定用户的兴趣得分，为用户分配兴趣标签的方式，能够有效的确定用户的兴趣主题，减少人工处理时间。

应清楚地理解，本发明描述了如何形成和使用特定示例，但本发明的原理不限于这些示例的任何细节。相反，基于本发明公开的内容的教导，这些原理能够应用于许多其它实施例。

图 4 是根据另一示例性实施例示出的一种用于确定用户兴趣标签的方法的流程图。由于数据量较大，直接使用 FP-growth 等关联算法找频繁集时会遇到计算时间过长或者存储不够无法计算等问题，此处可考虑编写 map-reduce 利用数据仓库的分布式计算架构实现此方法。图 4 是对由分词数据获取种子数据的示例性描述。

如图 4 所示，在 S402 中，根据预定条件，获取所述分词数据中所有的组合数据。在

本实施例中，基于如下的考虑：3 个或者小于 3 个词不足以定位用户的兴趣爱好，过大（如超过 15）则用户此单用户兴趣复杂且会导致后面的计算量过大，可例如选取产品词大于 3 且小于 15 的订单产品词列表参与后续计算；对于每一单的产品词列表，得到词量大于 3 的所有组合（此步可例如通过 map-reduce 实现）。例：(便签纸，加厚纸杯，卷纸，复印纸，抽纸，记事本子)大于 3 的组合共有 $C_6^4 + C_6^5 + C_6^6 = 22$ 种组合结果。

在 S404 中，对每一种组合数据，根据其订单数量，确定所述组合数据的频繁集。可例如订单量大于预定阈值的产品组合为频繁集。

在 S406 中，对所述频繁集进行最大频繁集计算，获取种子数据。对上一步得到的频繁集进行计算得到最大频繁集，将最大频繁集中的数据作为种子数据。种子数据结果如图 5 所示。

根据本发明的用于确定用户兴趣标签的方法，通过频繁集获取种子数据，进而将此种子数据作为 LDA 计算输入的方式，可以得到质量较高的兴趣主体，减少人工处理时间。

在本公开的一种示例性实施例中，还包括：通过历史数据，获取用户购买数据，所述购买数据包括购买产品次数以及购买产品标识。

图 6、7 是根据一示例性实施例示出的一种用于确定用户兴趣标签的方法的示意图。

在本公开的一种示例性实施例中，所述通过所述词向量数据与所述词权重数据确定用户的兴趣标签，包括：通过所述用户购买数据，确定所述用户的词向量数据以及词权重数据；通过所述用户的词向量数据以及词权重数据，计算所述用户的兴趣数值；通过所述兴趣数值确定所述用户的所述兴趣标签。将每一个最大频繁集作为 LDA 主题模型的种子词进行训练得到该兴趣下较为完整的词向量及每个词的权重。如图 6 所示（主题+词+词权重）。计算所有用户在一段时间内购买过的产品及每个产品的购买次数（用户账号+产品词+产品购买次数），结果如图 7 所示。

图 8、9 是根据一示例性实施例示出的一种用于确定用户兴趣标签的方法的示意图。

在本公开的一种示例性实施例中，所述通过所述用户的词向量数据以及词权重数据，计算所述用户的兴趣数值，包括：

$Sum = (a * Q)$;其中，Sum 为用户的所述兴趣数值，a 为用户购买产品次数，Q 为产品对应的词权重。还包括：判断所述兴趣数值是否大于预定阈值；以及将大于预定阈值的所述兴趣数值对应的兴趣标签确定为所述用户的兴趣标签。对于每一个用户，能够得到其每一个产品词所属的兴趣及产品词权重。如下图所示，能够得到用户 4 在园艺下的所有产品词及产品词权重，可例如，sum（产品购买次数*产品词权重）即为其园艺兴趣得分。得分情况如图 8 所示。当用户的兴趣得分大于某个阈值时，给用户打上相应的兴趣标签，结果如图 9 所示（主题、账号）。

在本公开的一种示例性实施例中，还包括：通过所述用户的所述兴趣标签进行信息推广。

图 10 是根据另一示例性实施例示出的一种用于确定用户兴趣标签的方法的流程图。

在 S1002 中，加工用户的购买数据。

在 S1004 中，获取订单产品词列表。

在 S1006 中，识别最大频繁集，确定种子词。

在 S1008 中，将种子词作为 LDA 的参数，得到兴趣此两项和词权重。

在 S1010 中，计算用户的产品词向量及产品的购买次数。

在 S1012 中，计算用户在每个兴趣上的得分，得到用户的兴趣标签。

获取用户在电商网站上的购物数据，首先使用频繁集的方法初步定位用户兴趣，得到种子词，再将种子词作为 LDA 的输入，得到能够比较全面刻画兴趣的产品词向量。对比兴趣的产品词向量和用户的产品词向量，对满足一定条件的用户打上相应的兴趣标签。

本领域技术人员可以理解实现上述实施例的全部或部分步骤被实现为由 CPU 执行的计算机程序。在该计算机程序被 CPU 执行时，执行本发明提供的上述方法所限定的上述功能。所述的程序可以存储于一种计算机可读存储介质中，该存储介质可以是只读存储器，磁盘或光盘等。

此外，需要注意的是，上述附图仅是根据本发明示例性实施例的方法所包括的处理的示意性说明，而不是限制目的。易于理解，上述附图所示的处理并不表明或限制这些处理的时间顺序。另外，也易于理解，这些处理可以是例如在多个模块中同步或异步执行的。

下述为本发明装置实施例，可以用于执行本发明方法实施例。对于本发明装置实施例中未披露的细节，请参照本发明方法实施例。

图 11 是根据一示例性实施例示出的一种用于确定用户兴趣标签的装置的框图。

基础模块 1102 用于将基础数据进行预处理，获取分词数据。

种子模块 1104 用于对所述分词数据进行最大频繁集识别，获取种子数据。

训练模块 1106 用于将所述种子数据进行数据训练，获取词向量数据与词权重数据。

标签模块 1108 用于通过所述词向量数据与所述词权重数据确定用户兴趣标签。

根据本发明的用于确定用户兴趣标签的装置，通过对原始数据进行分词表示，进而采用三层贝叶斯网络对分词数据进行训练，获得词向量以及词权重，进而确定用户的兴趣得分，为用户分配兴趣标签的方式，能够有效的确定用户的兴趣主题，减少人工处理时间。

图 12 是根据一示例性实施例示出的一种电子设备的框图。

下面参照图 12 来描述根据本发明的这种实施方式的电子设备 200。图 12 显示的电子设备 200 仅仅是一个示例，不应对本发明实施例的功能和使用范围带来任何限制。

如图 12 所示，电子设备 200 以通用计算设备的形式表现。电子设备 200 的组件可以包括但不限于：至少一个处理单元 210、至少一个存储单元 220、连接不同系统组件（包括存储单元 220 和处理单元 210）的总线 230、显示单元 240 等。

其中，所述存储单元存储有程序代码，所述程序代码可以被所述处理单元 210 执行，

使得所述处理单元 210 执行本说明书上述电子处方流转处理方法部分中描述的根据本发明各种示例性实施方式的步骤。例如，所述处理单元 210 可以执行如图 2，图 4 中所示的步骤。

所述存储单元 220 可以包括易失性存储单元形式的可读介质，例如随机存取存储单元（RAM）2201 和/或高速缓存存储单元 2202，还可以进一步包括只读存储单元（ROM）2203。

所述存储单元 220 还可以包括具有一组（至少一个）程序模块 2205 的程序/实用工具 2204，这样的程序模块 2205 包括但不限于：操作系统、一个或者多个应用程序、其它程序模块以及程序数据，这些示例中的每一个或某种组合中可能包括网络环境的实现。

总线 230 可以为表示几类总线结构中的一种或多种，包括存储单元总线或者存储单元控制器、外围总线、图形加速端口、处理单元或者使用多种总线结构中的任意总线结构的局域总线。

电子设备 200 也可以与一个或多个外部设备 300（例如键盘、指向设备、蓝牙设备等）通信，还可与一个或者多个使得用户能与该电子设备 200 交互的设备通信，和/或与使得该电子设备 200 能与一个或多个其它计算设备进行通信的任何设备（例如路由器、调制解调器等等）通信。这种通信可以通过输入/输出（I/O）接口 250 进行。并且，电子设备 200 还可以通过网络适配器 260 与一个或者多个网络（例如局域网（LAN），广域网（WAN）和/或公共网络，例如因特网）通信。网络适配器 260 可以通过总线 230 与电子设备 200 的其它模块通信。应当明白，尽管图中未示出，可以结合电子设备 200 使用其它硬件和/或软件模块，包括但不限于：微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、RAID 系统、磁带驱动器以及数据备份存储系统等。

通过以上的实施方式的描述，本领域的技术人员易于理解，这里描述的示例实施方式可以通过软件实现，也可以通过软件结合必要的硬件的方式来实现。因此，根据本公开实施方式的技术方案可以以软件产品的形式体现出来，该软件产品可以存储在一个非易失性存储介质（可以是 CD-ROM，U 盘，移动硬盘等）中或网络上，包括若干指令以使得一台计算设备（可以是个人计算机、服务器、或者网络设备等）执行根据本公开实施方式的上述电子处方流转处理方法。

图 13 是根据一示例性实施例示出的一种计算机可读介质示意图。

参考图 13 所示，描述了根据本发明的实施方式的用于实现上述方法的程序产品 400，其可以采用便携式紧凑盘只读存储器(CD-ROM)并包括程序代码，并可以在终端设备，例如个人电脑上运行。然而，本发明的程序产品不限于此，在本文件中，可读存储介质可以是任何包含或存储程序的有形介质，该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。

所述程序产品可以采用一个或多个可读介质的任意组合。可读介质可以是可读信号介质或者可读存储介质。可读存储介质例如可以为但不限于电、磁、光、电磁、红外线、或

半导体的系统、装置或器件，或者任意以上的组合。可读存储介质的更具体的例子（非穷举的列表）包括：具有一个或多个导线的电连接、便携式盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、光纤、便携式紧凑盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。

所述计算机可读存储介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了可读程序代码。这种传播的数据信号可以采用多种形式，包括但不限于电磁信号、光信号或上述的任意合适的组合。可读存储介质还可以是可读存储介质以外的任何可读介质，该可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。可读存储介质上包含的程序代码可以用任何适当的介质传输，包括但不限于无线、有线、光缆、RF 等等，或者上述的任意合适的组合。

可以以一种或多种程序设计语言的任意组合来编写用于执行本发明操作的程序代码，所述程序设计语言包括面向对象的程序设计语言—诸如 Java、C++等，还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算设备上执行、部分地在用户设备上执行、作为一个独立的软件包执行、部分在用户计算设备上部分在远程计算设备上执行、或者完全在远程计算设备或服务器上执行。在涉及远程计算设备的情形中，远程计算设备可以通过任意种类的网络，包括局域网（LAN）或广域网（WAN），连接到用户计算设备，或者，可以连接到外部计算设备（例如利用因特网服务提供商来通过因特网连接）。

上述计算机可读介质承载有一个或者多个程序，当上述一个或者多个程序被一个该设备执行时，使得该计算机可读介质实现如下功能：将基础数据进行预处理，获取分词数据；对所述分词数据进行最大频繁集识别，获取种子数据；将所述种子数据进行数据训练，获取词向量数据与词权重数据；以及通过所述词向量数据与所述词权重数据确定用户兴趣标签。

本领域技术人员可以理解上述各模块可以按照实施例的描述分布于装置中，也可以进行相应变化唯一不同于本实施例的一个或多个装置中。上述实施例的模块可以合并为一个模块，也可以进一步拆分成多个子模块。

通过以上的实施例的描述，本领域的技术人员易于理解，这里描述的示例实施例可以通过软件实现，也可以通过软件结合必要的硬件的方式来实现。因此，根据本发明实施例的技术方案可以以软件产品的形式体现出来，该软件产品可以存储在一个非易失性存储介质（可以是 CD-ROM，U 盘，移动硬盘等）中或网络上，包括若干指令以使得一台计算设备（可以是个人计算机、服务器、移动终端、或者网络设备等）执行根据本发明实施例的方法。

此外，本说明书说明书附图所示出的结构、比例、大小等，均仅用以配合说明书所公开的内容，以供本领域技术人员了解与阅读，并非用以限定本公开可实施的限定条件，故不具技术上的实质意义，任何结构的修饰、比例关系的改变或大小的调整，在不影响本公

开所能产生的技术效果及所能实现的目的下，均应仍落在本公开所公开的技术内容得能涵盖的范围内。同时，本说明书中所引用的如“上”、“第一”、“第二”及“一”等的用语，也仅为便于叙述的明了，而非用以限定本公开可实施的范围，其相对关系的改变或调整，在无实质变更技术内容下，当也视为本发明可实施的范畴。

权利要求

1、一种用于确定用户兴趣标签的方法，其特征在于，包括：

将基础数据进行预处理，获取分词数据；

对所述分词数据进行最大频繁集识别，获取种子数据；

将所述种子数据进行数据训练，获取词向量数据与词权重数据；以及

通过所述词向量数据与所述词权重数据确定用户兴趣标签。

2、如权利要求 1 所述的方法，其特征在于，所述将基础数据进行预处理，获取分词数据，包括：

通过用户历史购物数据生成所述基础数据；以及

对所述基础数据进行分词处理，生成所述分词数据。

3、如权利要求 1 所述的方法，其特征在于，所述对所述分词数据进行最大频繁集识别，获取种子数据，包括：

根据预定条件，获取所述分词数据中所有的组合数据；

对每一种组合数据，根据其订单数量，确定所述组合数据的频繁集；

对所述频繁集进行最大频繁集计算，获取种子数据。

4、如权利要求 1 所述的方法，其特征在于，所述对所述分词数据进行最大频繁集识别，获取种子数据，包括：

通过数据仓库的分布式计算架构，对所述分词数据进行最大频繁集识别，获取所述种子数据。

5、如权利要求 1 所述的方法，其特征在于，所述将所述种子数据进行数据训练，包括：

通过三层贝叶斯模型对所述种子数据进行数据训练。

6、如权利要求 1 所述的方法，其特征在于，还包括：

通过历史数据，获取用户购买数据，所述购买数据包括购买产品次数以及购买产品标识。

7、如权利要求 6 所述的方法，其特征在于，所述通过所述词向量数据与所述词权重数据确定用户的兴趣标签，包括：

通过所述用户购买数据，确定所述用户的词向量数据以及词权重数据；

通过所述用户的词向量数据以及词权重数据，计算所述用户的兴趣数值；

通过所述兴趣数值确定所述用户的所述兴趣标签。

8、如权利要求 7 所述的方法，其特征在于，所述通过所述用户的词向量数据以及词权重数据，计算所述用户的兴趣数值，包括：

$Sum = (a * Q)$ ；

其中，Sum 为用户的所述兴趣数值，a 为用户购买产品次数，Q 为产品对应的词权重。

9、如权利要求 7 所述的方法，其特征在于，所述通过所述兴趣数值确定所述用户的所述兴趣标签，还包括：

判断所述兴趣数值是否大于预定阈值；以及

将大于预定阈值的所述兴趣数值对应的兴趣标签确定为所述用户的兴趣标签。

10、如权利要求 1 所述的方法，其特征在于，还包括：

通过所述用户的所述兴趣标签进行信息推广。

11、一种用于确定用户兴趣标签的装置，其特征在于，包括：

基础模块，用于将基础数据进行预处理，获取分词数据；

种子模块，用于对所述分词数据进行最大频繁集识别，获取种子数据；

训练模块，用于将所述种子数据进行数据训练，获取词向量数据与词权重数据；以及

标签模块，用于通过所述词向量数据与所述词权重数据确定用户兴趣标签。

12、一种电子设备，其特征在于，包括：

一个或多个处理器；

存储装置，用于存储一个或多个程序；

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如权利要求 1-10 中任一所述的方法。

13、一种计算机可读介质，其上存储有计算机程序，其特征在于，所述程序被处理器执行时实现如权利要求 1-10 中任一所述的方法。

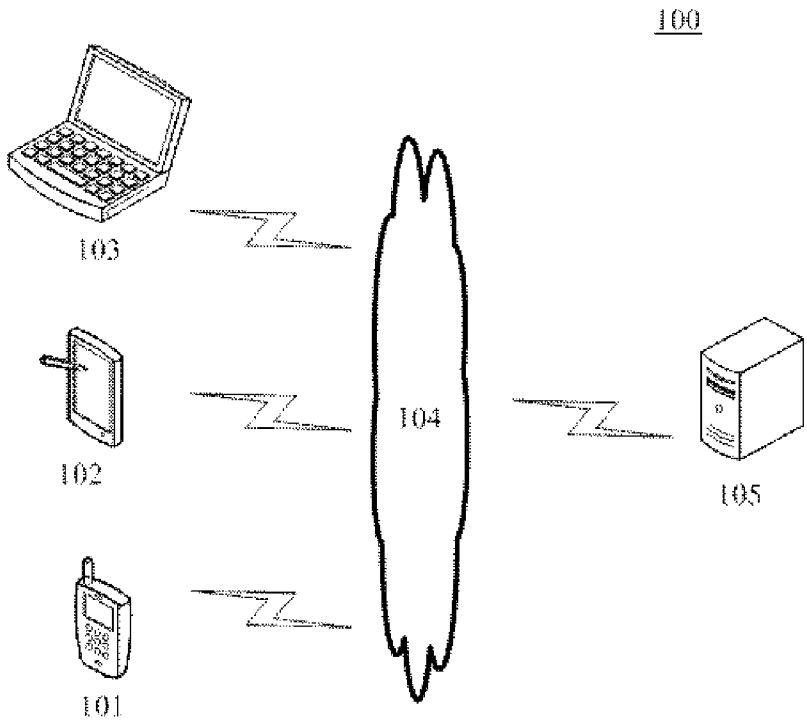
100
103
102
101
104
105

图 1

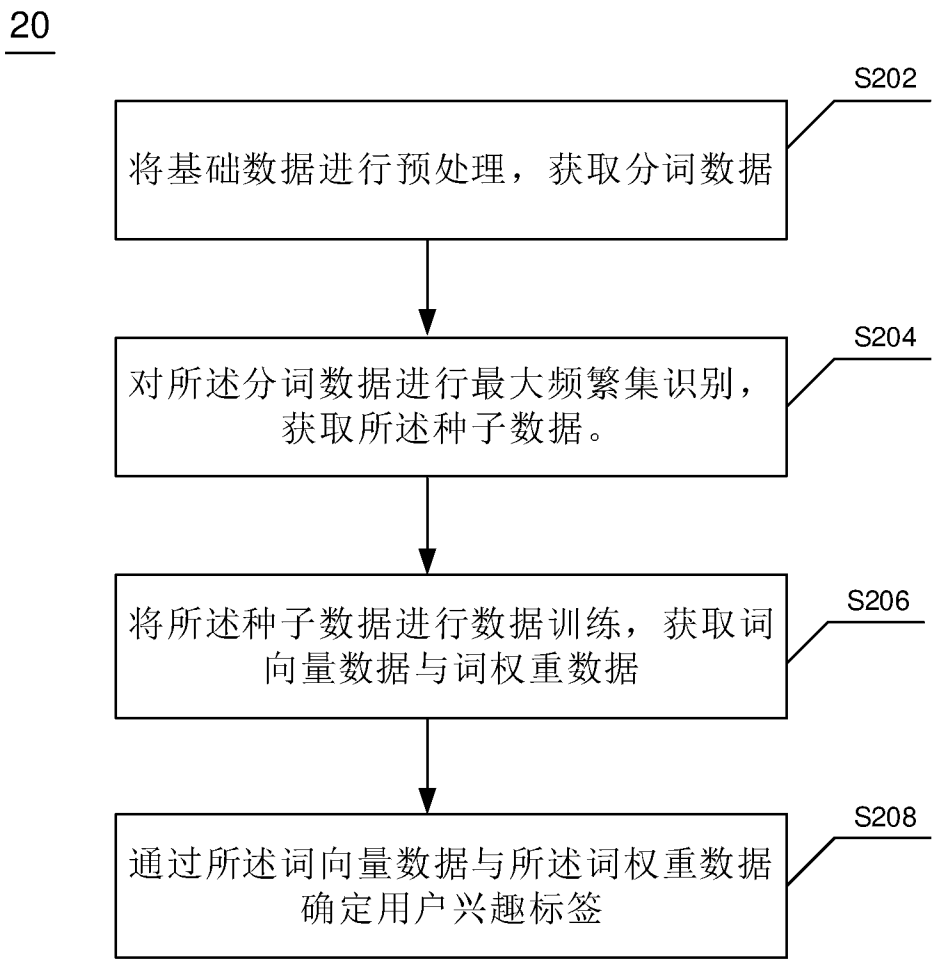
20
S202
将基础数据进行预处理，获取分词数据
S204
对所述分词数据进行最大频繁集识别，获取所述种子数据。
S206
将所述种子数据进行数据训练，获取词向量数据与词权重数据
S208
通过所述词向量数据与所述词权重数据确定用户兴趣标签

图 2

用户账号	订单	商品id	商品名
用户1	订单1	1150546	百草味 水果干 芒果干120g/袋 菲律宾风味 蜜饯零食小吃干果果脯
用户1	订单1	1611790	卫龙 大面筋（香辣味）68g
用户1	订单1	1203730	韩国进口 zek 芝士鳕鱼肠 105g
用户1	订单1	1690511	良品铺子 芒果味麻薯 爆浆麻薯 糕点点心 零食特产小吃150g
用户1	订单1	1278110	百草味 海鲜零食特产 风琴鱿鱼片80g/袋 休闲小吃 手撕鱿鱼丝
用户2	订单2	1637286	苏泊尔 SUPOR 312170-01-LS 304不锈钢 毛巾架 浴巾架浴室挂件 套装
用户2	订单2	1028426084	瑞士希能/HOROW 不锈钢浴室三角篮卫生间双层角篮置物架三角架
用户2	订单2	2188554	名爵 (MEJUE) Z-3101不锈钢厨房置物架 厨房挂件多功能收纳架 挂架刀架 调料架 壁挂 双杯置物架
用户3	订单3	2848246	派滋 魅蓝3s手机壳/魅族魅蓝3手机壳硅胶防摔保护套 透明
用户4	订单3	3155906	魅族 魅蓝3s三个月碎屏换新（套装版）
用户5	订单3	1336842	魅族 EP-21耳机
用户6	订单3	2908677	【超值套装版】魅族 魅蓝3S 全网通公开版 16GB 银色 移动联通电信4G手机 双卡双待

订单	产品词列表
订单4	便签纸,加厚纸杯,卷纸,复印纸,抽纸,记事本子
订单5	导出液,懒人霸,眉粉,美容仪器
订单6	千岛酱,口香糖,意大利面,挂面,沙司,油咖喱,泡菜,火腿肠,火锅,火锅底料,粉丝,粥米,腐竹,芝麻酱,黑米,黑胡椒酱
订单7	仙贝,坚果,果干,红包,蛋饼,饼干
订单8	年糕条,西饼,饼干
订单9	保温杯,卫生巾,多用剪,玻璃水杯,电子秤
订单10	低筋面粉,卫生纸,方便面,果干,桂圆干,火腿肠,红枣,腌料,面粉
订单11	抽纸,放大器,洁厕液,胶带
订单12	卷纸,坚果,果干,洗衣液,纯牛奶
订单13	发蜡,染发霜,牙刷,牙膏

图 3

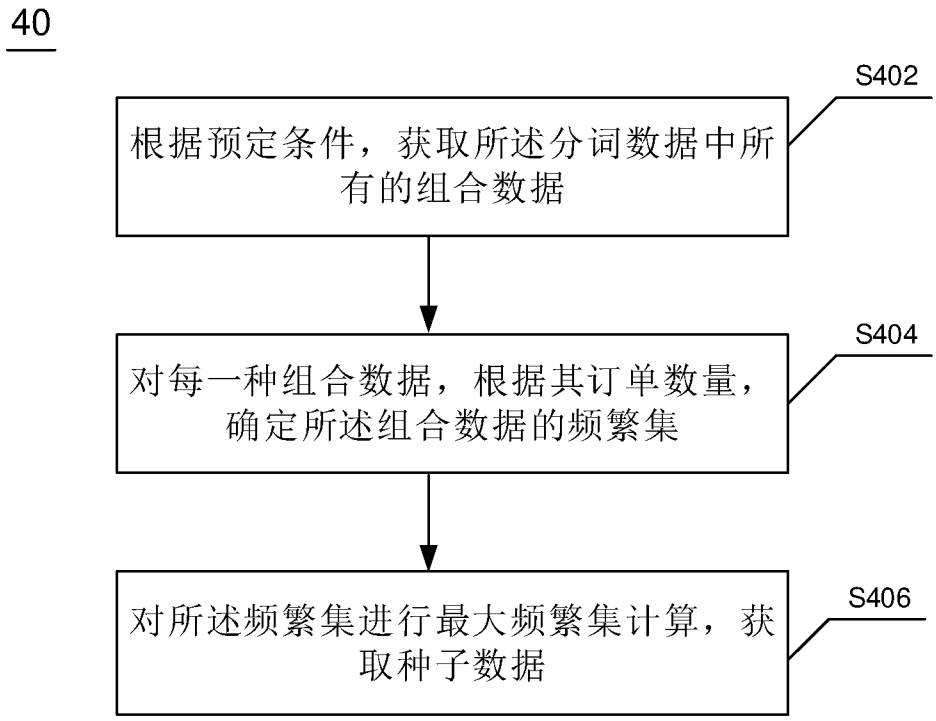

图 4

读卡器,镜头布,单肩包,纸巾,存储卡,录机,擦镜纸,微单包,相机包,传感器,除尘纸,微单
硅胶垫,裱花枪,纸托,饼干,裱花器,烘焙工具,蛋糕模具,裱面垫,脱模刀,刮刀
文胸,瑜伽服,瑜伽垫,网袋,健身服
足浴粉,木桶,泡脚桶,木盆,按摩珠,火山石
蕾莲盆,种球,营养土,睡莲盆,营养液
海洋球,洗头帽,浴网,玩具,水温计,浴桶,沐浴盆,浴盆,水勺

图 5

园艺	吊兰	0.003544945
园艺	多菌灵	0.004252519
园艺	肥料	0.02485646
园艺	分装瓶	0.003670248
园艺	复合肥	0.007033806
园艺	富贵竹	0.004252519
园艺	果苗	0.00368646
园艺	果树苗	0.012870219
园艺	花卉	0.004252519
园艺	花架	0.014466754
园艺	花架子	0.009678791
园艺	花苗	0.012093029
园艺	花盆	0.04263924
园艺	花盆架	0.010882807
园艺	花瓶	0.011903155
园艺	花洒	0.006979651
园艺	花束	0.003261916
园艺	花种	0.00936994
园艺	浇花壶	0.004252519
园艺	绿萝	0.010719395
园艺	玫瑰花苗	0.007998
园艺	泥炭土	0.003969489
园艺	喷壶	0.016779348
园艺	喷水壶	0.009065013
园艺	喷雾瓶	0.003630376
园艺	喷雾器	0.008891472
园艺	盆景	0.014357265
园艺	盆栽	0.08173318
园艺	洒水壶	0.009975872
园艺	生根粉	0.01294604

图 6

用户账号	产品词	产品购买次数
用户4	草莓苗	1
用户4	果苗	1
用户4	花架子	1
用户4	喷壶	2
用户4	喷水壶	2
用户4	吊兰	1
用户4	多菌灵	1
用户5	加湿器	1
用户5	玩具	1
用户5	连体衣	2
用户5	连衣裙	1

图 7

用户账号	产品词	产品购买次数	产品词权重	兴趣得分
用户4	草莓苗	1	0.00340343	
用户4	果苗	1	0.00368646	0.00340343*1+0.0036864
用户4	花架子	1	0.009678791	6*1+0.009678791*1+0.01
用户4	喷壶	2	0.016779348	6779348*2+0.009065013*
用户4	喷水壶	2	0.009065013	2+0.003544945*1+0.0042
用户4	吊兰	1	0.003544945	52519*1=0.076255
用户4	多菌灵	1	0.004252519	

图 8

图 9

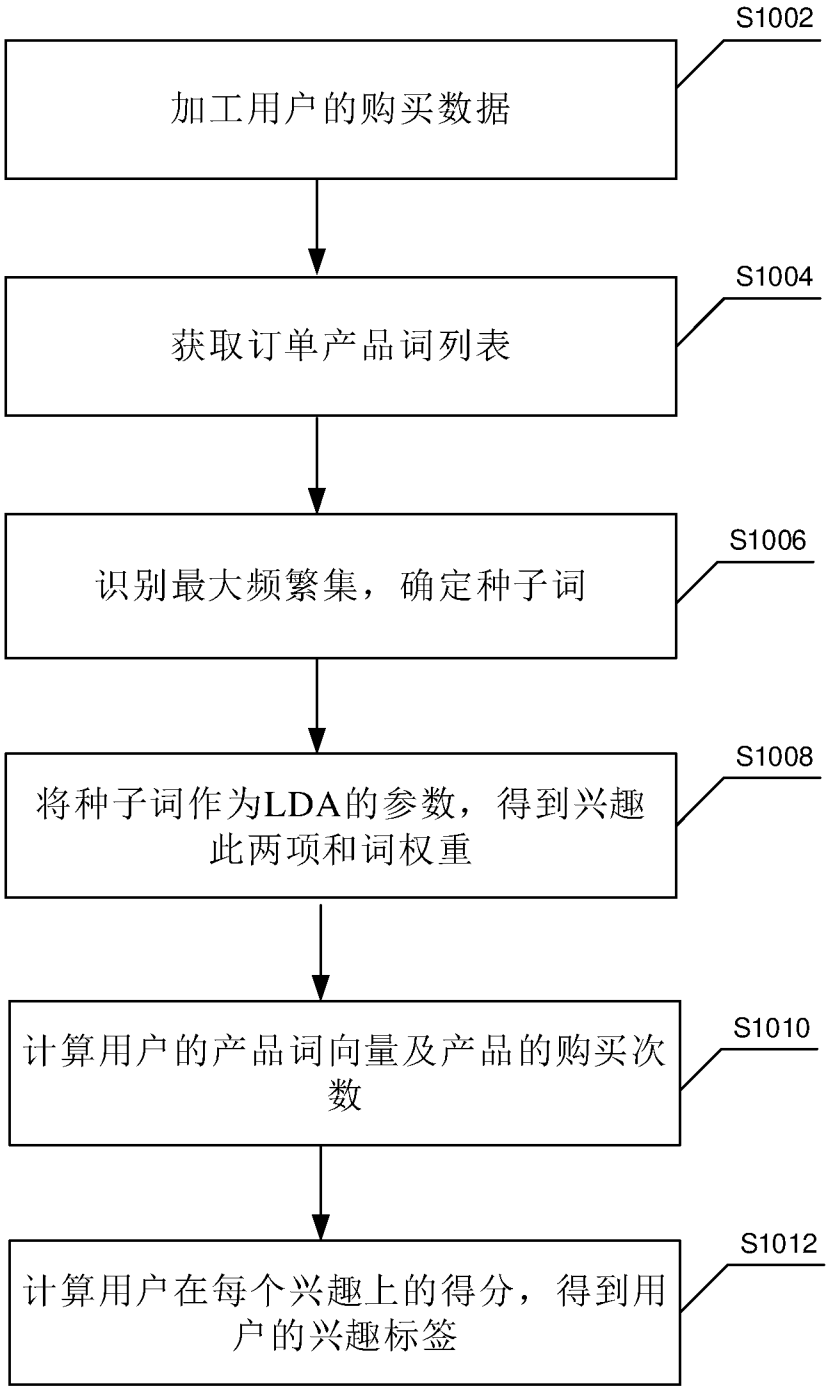
S1002
加工用户的购买数据
S1004
获取订单产品词列表
S1006
识别最大频繁集，确定种子词
S1008
将种子词作为LDA的参数，得到兴趣此两项和词权重
S1010
计算用户的产品词向量及产品的购买次数
S1012
计算用户在每个兴趣上的得分，得到用户的兴趣标签

图 10

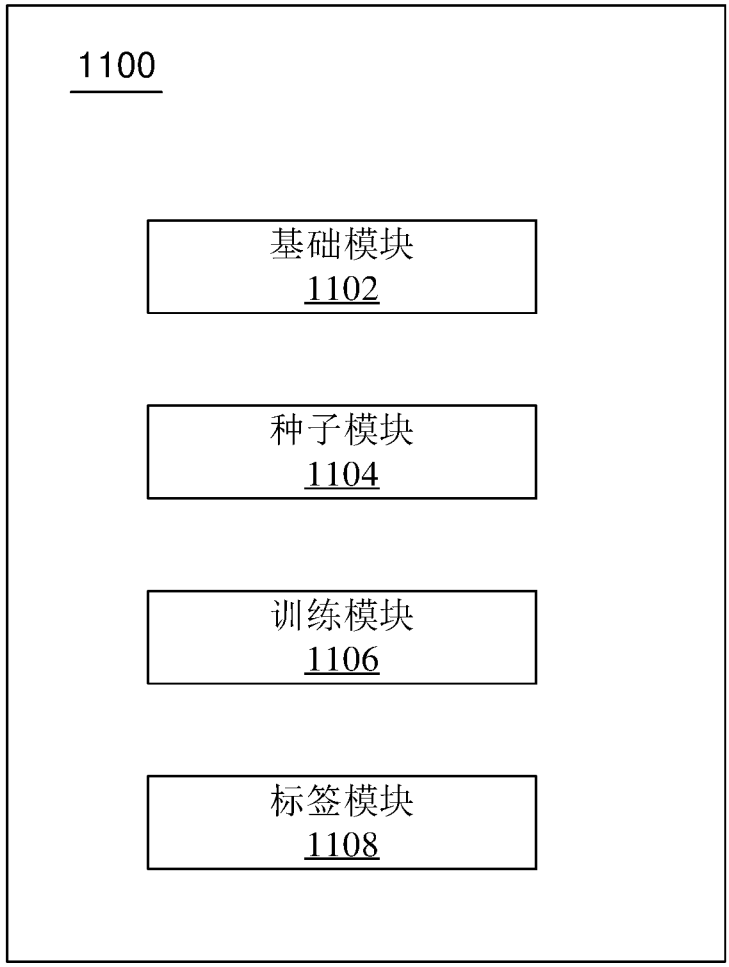
1100
基础模块
1102
种子模块
1104
训练模块
1106
标签模块
1108

图 11

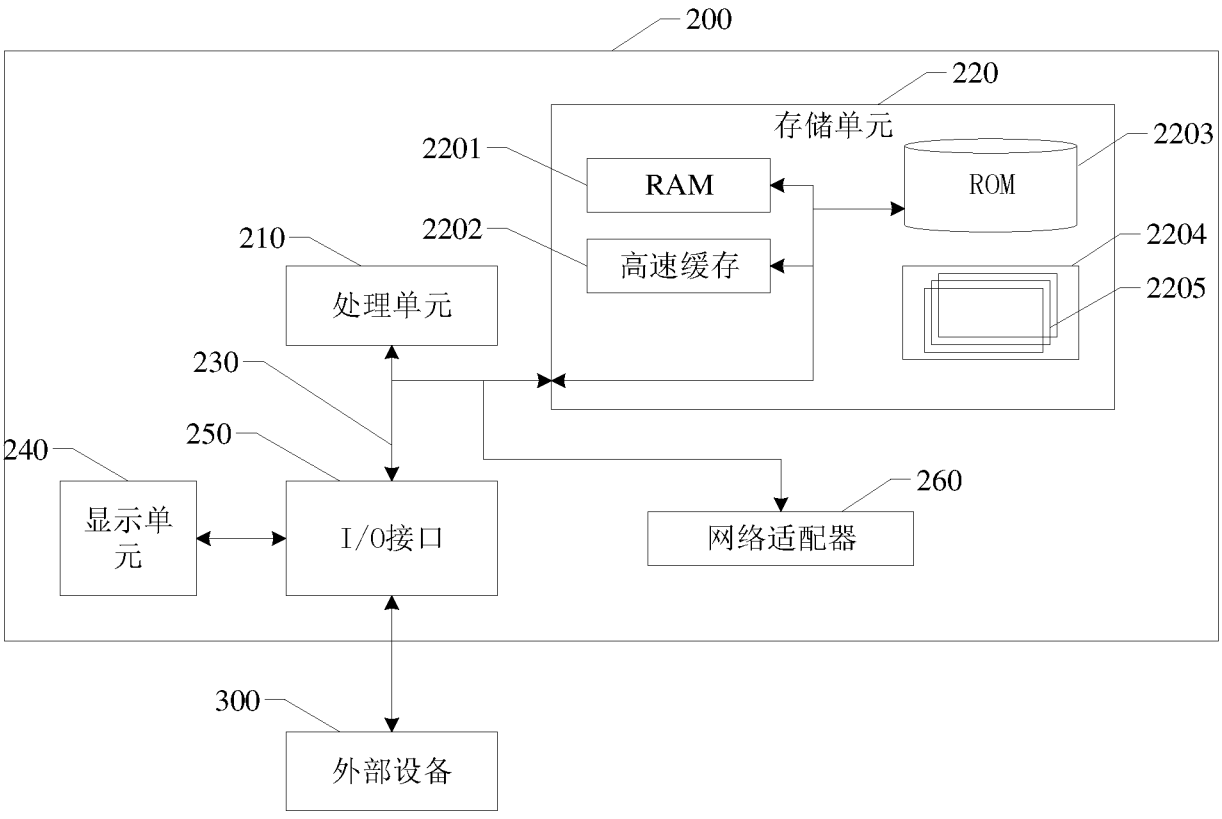
200
220
存储单元
2201
RAM
ROM
2203
2202
高速缓存
2204
2205
210
处理单元
230
250
240
260
显示单元
I/O接口
网络适配器
300
外部设备

图 12

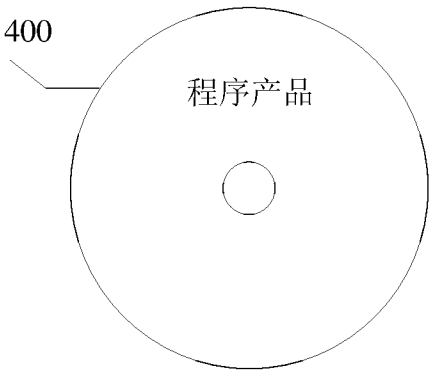

图 13

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/107969

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/62(2006.01)i; G06F 17/30(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06Q; G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, WPI, EPODOC, CNKI: 分词, 频繁集, 种子, 训练, 向量, 矢量, 权重, 兴趣, 偏好, 标签, 三层贝叶斯, 购买, 订单, 行为, 营销画像, 次数, 频率, 支持度, 狄利克雷分配, LDA, latent, dirichlet, allocation, maximal, frequent, itemset, seed, word, segmentation, training, vector, weight, interest, label, order, history

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 106649681 A (BEIJING KINGSOFT SECURITY SOFTWARE CO., LTD.) 10 May 2017 (2017-05-10) abstract, and description, paragraphs [0002]-[0013] and [0050]	1-13
Y	CN 101206752 A (BEIJING KEWEN SHUYE INFORMATION TECHNOLOGY CO., LTD.) 25 June 2008 (2008-06-25) abstract, and description, pages 1-3	1-13
PX	CN 107729937 A (BEIJING JINGDONG SHANGKE INFORMATION TECHNOLOGY CO., LTD. ET AL.) 23 February 2018 (2018-02-23) claims 1-13	1-13
A	CN 101122909 A (HITACHI LTD. ET AL.) 13 February 2008 (2008-02-13) entire document	1-13
A	CN 105677769 A (GUANGZHOU SHENMA MOBILE INFORMATION TECHNOLOGY CO., LTD.) 15 June 2016 (2016-06-15) entire document	1-13

☑ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 November 2018	**04 January 2019**

Name and mailing address of the ISA/CN	Authorized officer
State Intellectual Property Office of the P. R. China **No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing** **100088** **China**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2018/107969

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 103593400 A (SHAANXI PROVINCIAL METEOROLOGICAL BUREAU) 19 February 2014 (2014-02-19) entire document	1-13

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2018/107969

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 106649681 A	10 May 2017	None	
CN 101206752 A	25 June 2008	None	
CN 107729937 A	23 February 2018	None	
CN 101122909 A	13 February 2008	None	
CN 105677769 A	15 June 2016	WO 2017114019 A1	06 July 2017
		US 2018307680 A1	25 October 2018
CN 103593400 A	19 February 2014	None	

国际检索报告

国际申请号 PCT/CN2018/107969

A. 主题的分类

G06K 9/62(2006.01)i; G06F 17/30(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06Q; G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用）)

CNPAT, WPI, EPODOC, CNKI: 分词, 频繁集, 种子, 训练, 向量, 矢量, 权重, 兴趣, 偏好, 标签, 三层贝叶斯, 购买, 订单, 行为, 营销画像, 次数, 频率, 支持度, 狄利克雷分配, LDA, latent, dirichlet, allocation, maximal, frequent, itemset, seed, word, segmentation, training, vector, weight, interest, label, order, history

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
Y	CN 106649681 A (北京金山安全软件有限公司) 2017年 5月 10日 (2017-05-10) 摘要，说明书第[0002]-[0013]、[0050]段	1-13
Y	CN 101206752 A (北京科文书业信息技术有限公司) 2008年 6月 25日 (2008-06-25) 摘要，说明书第1-3页	1-13
PX	CN 107729937 A (北京京东尚科信息技术有限公司 等) 2018年 2月 23日 (2018-02-23) 权利要求1-13	1-13
A	CN 101122909 A (株式会社日立制作所 等) 2008年 2月 13日 (2008-02-13) 全文	1-13
A	CN 105677769 A (广州神马移动信息科技有限公司) 2016年 6月 15日 (2016-06-15) 全文	1-13
A	CN 103593400 A (陕西省气象局) 2014年 2月 19日 (2014-02-19) 全文	1-13

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

"A" 认为不特别相关的表示了现有技术一般状态的文件

"E" 在国际申请日的当天或之后公布的在先申请或专利

"L" 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

"O" 涉及口头公开、使用、展览或其他方式公开的文件

"P" 公布日先于国际申请日但迟于所要求的优先权日的文件

"T" 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

"X" 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

"Y" 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

"&" 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2018年 11月 27日	2019年 1月 4日
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 费聿辉 电话号码 86-(10)-53961778

表 PCT/ISA/210 (第2页) (2015年1月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2018/107969

检索报告引用的专利文件			公布日（年/月/日）	同族专利			公布日（年/月/日）
CN	106649681	A	2017年 5月 10日	无			
CN	101206752	A	2008年 6月 25日	无			
CN	107729937	A	2018年 2月 23日	无			
CN	101122909	A	2008年 2月 13日	无			
CN	105677769	A	2016年 6月 15日	WO	2017114019	A1	2017年 7月 6日
				US	2018307680	A1	2018年 10月 25日
CN	103593400	A	2014年 2月 19日	无			